Recovery of motion parameters from distortions in scanned images

Jeffrey B. Mulligan
NASA Ames Research Center

Abstract: Scanned images, such as those produced by the scanning-laser ophthalmoscope (SLO), show distortions when there is target motion. This is because pixels corresponding to different image regions are acquired sequentially, and so, in essence, are slices of different snapshots. While these distortions create problems for image registration algorithms, they are potentially useful for recovering target motion parameters at temporal frequencies above the frame rate. Stetter, Sendtner and Timberlake (Vision Res, vol. **36**, pp. 1987-1994, 1996) measured large distortions in SLO images to recover the time course of rapid horizontal saccadic eye movements. Here, this work is extended with the goal of automatically recovering small eye movements in two dimensions. Eye position during the frame interval is modeled using a low dimensional parametric description, which in turn is used to generate predicted distortions of a reference template. The input image is then registered to the distorted template using normalized cross correlation. The motion parameters are then varied, and the correlation recomputed, to find the motion which maximizes the peak value of the correlation. The location and value of the correlation maximum are determined with sub-pixel precision using biquadratic interpolation, yielding eye position resolution better than 1 arc minute (Mulligan, Behavior Research Methods, Instruments and Computers, vol. **29**, pp. 54-65, 1997). This method of motion parameter estimation is tested using actual SLO images as well as simulated images. Motion parameter estimation might also be applied to individual video lines in order to reduce pipeline delays for a near real-time system.

1. Introduction

Video image sequences are often used to track object motion. Unless a special high frame-rate camera is used, the recovered motion is usually sampled in time at the video frame rate (50-60 Hz). While low resolution sampling is adequate for many applications, documentation of high-speed events often requires higher temporal resolution. For images obtained with a scanned system, in which individual pixel values are acquired at different times, it is possible to obtain higher temporal resolution for the motion of extended targets. The sequential nature of the scanning process introduces geometric distortions in the image of a moving target. By measuring these distortions, high temporal resolution information about the target motion can be recovered. This technique is especially useful when *a priori* knowledge about the possible target motions permits a concise description using low-dimensional parametric models, because this reduces the space of possible distortions which must be searched. In the following sections, expressions for the precise form of the motion-induced distortions will be derived.

Address for author correspondence: MS 262-2, NASA Ames Research Center, Moffett Field, CA, 94035-1000. Email: jbm@vision.arc.nasa.gov. An electronic version of this paper with additional multimedia figures is available at: http://vision.arc.nasa.gov/~jbm/papers/irw97.html.

1.1. Raster scanning

x, y	position in image plane
$s_x(t), s_y(t)$	position of scan at time t
f_L	line frequency (~15 kHz)
f_F	frame rate (~60 Hz)
i_L	index of current line
t_S	start time of current line
t_L	time in current line, $t - t_S$
$v_{S,x}, v_{S,y}$	scan velocities

Some imaging systems, using an electronic or mechanical shutter, can simultaneously capture all of the pixels in an image. In a scanned system, however, only a single point is sensed at a given time, and the location of this point is swept over the image area by electronic or mechanical means. Here we present some definitions and conventions that will allow us to precisely describe the scanning process.

The imaging area is defined to be a rectangle indexed by normal Cartesian coordinates x and y. The raster is defined by two scan functions, $s_x(t)$ and $s_y(t)$, which represent the instantaneous beam position. These functions are approximated by sawtooth waveforms (see figure 1). By convention, the horizontal dimension is scanned at a relatively high frequency, called the *line frequency*, f_L, while the slower vertical frequency determines the *frame rate*, f_F.

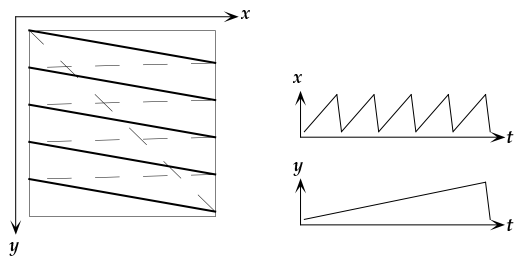

Figure 1: Diagram of raster pattern on the left, with the active portion of each line shown as a heavy solid line, and retrace as a dashed line (see appendix). On the right, the scan functions are shown over time.

Time $t = 0$ in our temporal coordinate system is the beginning of the current frame. By convention, numbering of raster lines begins with 1; The index of the current line, i_L, is

$$i_L = \lfloor t \; f_L \rfloor \; . \qquad (1)$$

We define t_S to be the time of the start of the current line, and t_L to be the time relative to the start of the current line:

$$t_S = \frac{i_L}{f_L} \; , \quad \text{and} \quad t_L = t - t_S . \qquad (2a,b)$$

These quantities are illustrated graphically in figure 2.

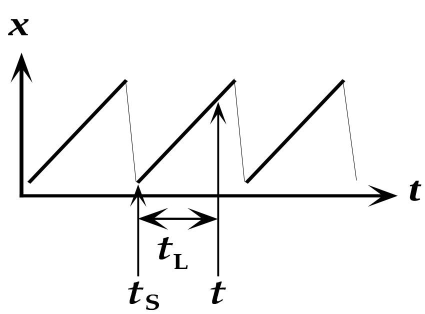

Figure 2: Raster waveform diagram, indicating the current time, t, the start time of the current line, t_S, and the time within the current line, t_L.

The *scan velocities*, $v_{S,x}$ and $v_{S,y}$, describe the rate at which the scanning beam traverses the image plane. When expressed in units of image widths per second, these are approximately equal to the scan frequencies, f_L and f_F (see appendix for details). We can write simple expressions for the instantaneous scan position in terms of the scan velocities. The horizontal scan position $s_x(t)$ is:

$$s_x(t) = t_L \; v_{S,x} \; , \qquad (3a)$$

In most scanning systems, the vertical scan is continuous (partly due to "mechanical" constraints), and

$$s_y(t) = t \; v_{S,y} \; . \qquad (3b)$$

1.2. Effects of object motion

P	a target point
$p_x(t), p_y(t)$	instantaneous position of P
$\dot{p}_x(t), \dot{p}_y(t)$	instantaneous velocity of P
$\ddot{p}_x(t), \ddot{p}_y(t)$	instantaneous acceleration P
x_0, y_0	position of P at time t=0
x_P, y_P	position of P in scanned image
t_P	time P is scanned

We consider a fiducial point on the target, located at coordinates (x_0, y_0) at time 0. Let the position at time t be expressed by the functions $p_x(t)$ and $p_y(t)$. These positions can be expressed using Taylor series, where $\dot{p}_x(0)$ is the x velocity at time 0, $\ddot{p}_x(0)$ is the acceleration, and so on:

$$p_x(t) = x_0 + \dot{p}_x(0)\; t + \tfrac{1}{2}\, \ddot{p}_x(0)\; t^2 + \cdots \qquad (4a)$$

$$p_y(t) = y_0 + \dot{p}_y(0)\; t + \tfrac{1}{2}\, \ddot{p}_y(0)\; t^2 + \cdots \qquad (4b)$$

We wish to know the position of the given point, (x_P, y_P), in the acquired image. When the point's trajectory intersects the raster, the time at which the point is scanned, t_P, will be:

$$t_P = \frac{y_P}{v_{S,y}} + \frac{x_P}{v_{S,x}} \; . \qquad (5a)$$

$$\approx \frac{y_P}{v_{S,y}} \; . \qquad (5b)$$

By making the approximation, we ignore the dependence on horizontal position. This is justified on the grounds that $v_{S,x}$ is large, and so this term will be small. By definition, $y_P = p_y(t_P)$, and so the value of t_P obtained in equation 5b may be substituted into equation 4b, which can then be solved for y_P. The result can then be used to evaluate equation 4a to obtain x_P.

In general, the raster will not pass directly over the point, and features of finite size will often be represented in more than one scan line. We assume that little target motion occurs during a single line time, so the position of a feature located between two scan lines can be accurately determined by interpolation, and results obtained for points lying directly on the raster will hold for all points.

1.3. Example: constant object velocity

$v_{T,x}, v_{T,y}$	target velocity, $\dot{p}_x(t) = v_{T,x}$

We can use the results of the preceding section to generate simulated distorted images for various motions. We consider first the simple case where the target moves with constant velocity, $(v_{T,x}, v_{T,y})$:

$$p_x(t) = x_0 + v_{T,x} t \ , \tag{6a}$$

$$p_y(t) = y_0 + v_{T,y} t \ . \tag{6b}$$

Following the strategy outlined above, we construct the following equation for y_P:

$$y_P = y_0 + \frac{v_{T,y}\ y_P}{v_{S,y}} \ . \tag{7a}$$

$$= \frac{y_0\ v_{S,y}}{v_{S,y} - v_{T,y}} \ , \tag{7b}$$

$$= y_0 + \frac{v_{T,y}\ y_0}{v_{S,y} - v_{T,y}} \ . \tag{7c}$$

We can use this result to derive a corresponding expression for x_P:

$$x_P = x_0 + \frac{v_{T,x}\ y_0}{v_{S,y} - v_{T,y}} \ . \tag{8}$$

Several important points may be noted from these equations: first, the deviation in feature position for each component is proportional to y_0, the vertical position of the feature in the image, and to the corresponding component of object velocity. We also notice that when $v_{T,y} \geq v_{S,y}$ (object moving faster than the raster), the solution corresponds to a negative value of t, and does not correspond to a point in the current frame.

Distortions arising from a target speed of $\frac{v_{S,y}}{4}$ are illustrated in figure 3. The left-hand patch shows the image obtained when a square grid target is moved at the right, while the right-hand patch shows the image resulting from upward motion.

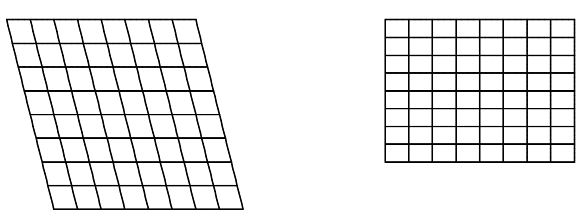

Figure 3: Image distortions of a regular grid for constant velocity motion to the right (left) and upwards (right).

1.4. Example: constant object acceleration

a_x, a_x	target acceleration, $\ddot{p}_x(t) = a_x$

We assume the object accelerates from rest at time 0 with accelerations a_x and a_y:

$$p_x(t) = x_0 + \tfrac{1}{2} a_x t^2 \ , \tag{9a}$$

$$p_y(t) = y_0 + \tfrac{1}{2} a_y t^2 \ . \tag{9b}$$

We first consider the case where $a_y = 0$, i.e. a purely horizontal motion. In this case, the vertical position of the fiducial point will not be changed, and the raster will scan the point at time $t_P = \frac{y_0}{v_{S,y}}$. Substituting this value into equation 9a, we obtain:

$$p_x(t_P) = x_0 + \frac{a_x\ p_y^2(0)}{2\ v_{S,y}^2} \ . \tag{10}$$

Equation 10 is quite similar to equation 8, except that here the deviation is proportional to the *square* of the vertical position. This case is illustrated on the left side of figure 4.

The case of vertical accelerations is more complex, due to the interaction between the accelerating motion with the vertical scan. As we did above with equation 7a, we begin by constructing an equation in y_P:

$$y_P = y_0 + \frac{a_y\ y_P^2}{2\ v_{S,y}^2} \ , \tag{11a}$$

which after application of the quadratic formula yields:

$$y_P = \frac{v_{S,y}\ (\ v_{S,y} \pm \gamma\ )}{a_y} \ , \tag{11b}$$

where

$$\gamma = \sqrt{v_{S,y}^2 - 2\, a_y\, y_0} \ . \tag{12}$$

The smaller of two solutions corresponds to the first coincidence of the raster and the point, while the larger only exists when the acceleration is so large that the point subsequently overtakes the raster. When the acceleration is so large that the raster *never* encounters the point, γ is imaginary.

After more algebra, we can also obtain this result for x_P:

$$x_P = x_0 + \frac{v_{S,y}^2 - a_y\, y_0 \pm v_{S,y}\, \gamma}{a_y} \ . \tag{13}$$